

GOOGLE SCHOLAR COMO FERRAMENTA DE ANÁLISE BIBLIOMÉTRICA EM BIBLIOTECAS ACADÊMICAS:

um estudo preliminar

Eixo temático: MÉTODOS, TÉCNICAS E FERRAMENTAS PARA ESTUDOS
BIBLIOMÉTRICOS E CIENTOMÉTRICOS

Modalidade: Apresentação oral

1 INTRODUÇÃO

Os desenvolvimentos tecnológicos, principalmente os da última década, têm levado a repensar os papéis de diversos tipos de instituições. No âmbito acadêmico, dispositivos agregadores de grande quantidade de tecnologia de informação e comunicação (smartphones, tablets, laptops e seus aplicativos, são os exemplos mais comuns), permitem a alunos, professores e pesquisadores o acesso a grandes quantidades de informação com maior mobilidade, comunicação instantânea e trabalho colaborativo e em Rede. Neste novo contexto, novos hábitos de acesso, leitura e memorização estão sendo desenvolvidos e estão alterando, inclusive, a forma e o processo de aprendizagem (por exemplo, através das plataformas de ensino online) (GARCIA, 2007). Conforme o 3º plano estratégico da REBIUN (Red Española de Bibliotecas Universitarias),

La evolución constante e imparable de las Tecnologías de la Información y Comunicación, la extensión social del uso de la red de Internet, los nuevos programas informáticos y soportes móviles que lanza el mercado, la irrupción y la importancia de las redes sociales están transformando todos los sistemas de formación superior, información y documentación como jamás se había visto hasta ahora (REBIUN, CRUE, 2011).

Cada vez mais, aplicativos (apps, api's) permitem a integração de sistemas¹, promovendo novos processos na publicação científica, centralizando na *Web* (Cloud) conteúdos e serviços associados, inclusive ferramentas interativas e acesso a conjuntos de dados digitais que do ponto de vista prático, não poderiam ser disponibilizados através de meios impressos.

Tradicionalmente as bibliotecas possuem práticas que envolvem regras, normas e códigos que atenderam relativamente bem ao desenvolvimento local de coleções impressas (físicas). Embora um número de bibliotecas participe ativamente do desenvolvimento e curadoria de Bibliotecas e Repositórios Digitais, a grande maioria assiste passivamente a estes importantes desenvolvimentos, além de já terem presenciado a migração dos Periódicos Científicos do impresso para o digital online. Embora os movimentos de Acesso Aberto tenham ganhado fôlego na última década, as bibliotecas acadêmicas convivem com Periódicos e agora livros sendo publicados eletronicamente e mantidos nas “bibliotecas digitais” dos Editores sob condições comerciais. As bibliotecas estão deixando de possuírem cópias próprias de seus periódicos, e o mesmo pode, muito em breve, acontecer com os livros, bastando conectar os usuários aos Editores. Ou seja, dentre as novas condições para as bibliotecas, em especial as acadêmicas, é o fato de que elas estão cada vez mais dependentes de documentos que não possuem, simplesmente fornecendo acesso a estes. Daí o ditado, "das coleções para as conexões" (NIELSEN; HJØRLAND, 2014).

Dentro deste contexto, a instituição biblioteca necessita estar ciente desta nova realidade. Além de inovar suas instalações, espaços e serviços e desenvolver projetos em colaboração com outras instituições, é preciso expandir suas ofertas de serviços para novas demandas. E o profissional bibliotecário necessita melhorar e renovar todas as suas habilidades, conhecimentos e competências. Em relação à pesquisa,

las bibliotecas deberán responder a las necesidades cada vez más exigentes de los investigadores. La estrategia va a ser trabajar junto a ellos y estar presente desde el inicio en los procesos de investigación que se llevan a cabo como proveedores de información científica, como editores de sus publicaciones y como promotores y difusores de sus resultados (REBIUN; CRUE, 2011).

¹Novas abordagens para o futuro da Publicação Científica: integração “Fiduswriter” e “Open Journal System”. <http://openaccess.sdum.uminho.pt/?p=3015>

Oferecer apenas acesso aos mais importantes periódicos científicos, catalogar materiais e colocá-los nas estantes para que sejam consultados pelos usuários e gerar simples estatísticas de uso do acervo local não é mais suficiente. As bibliotecas acadêmicas devem criar novos serviços bibliotecários personalizados para os pesquisadores, devem aprender a serem especialistas na gestão e exploração das métricas da ciência, de forma a apoiar e potencializar a nova comunicação científica, administrando a produção, edição e difusão das publicações. A visibilidade e o impacto da Universidade e de seus pesquisadores na sociedade em rede são demandas que as bibliotecas necessitam atender.

2 BIBLIOTECAS ACADÊMICAS E SERVIÇO DE DADOS BIBLIOMÉTRICOS

Em universidades, um dos aspectos básicos para a captação de financiamento de apoio à pesquisa é o seu resultado. Mais precisamente o impacto da pesquisa na sociedade. Estes resultados são divulgados principalmente por periódicos científicos, que são importantes meios de comunicação da ciência. Muitos departamentos de graduação, centros de pesquisa, laboratórios e programas de pós-graduação, possuem suas próprias revistas científicas que também são mantidas através de apoio. Desta forma, no atual contexto, torna-se ainda mais imprescindível a modernização e internacionalização dos veículos de divulgação científica. Indicadores bibliométricos são ferramentas essenciais para medir a produção e o impacto da pesquisa.

Bibliotecas acadêmicas, como instituições de base para a pesquisa, devem se preparar para fornecer serviços de dados bibliométricos.

3 Uso do GOOGLE SCHOLAR e o Software PUBLISH OR PERISH como ferramentas bibliométrias

Para explorar formas de obter dados e gerar índices métricos, optou-se por utilizar o Google Scholar (Google Acadêmico), uma ferramenta de pesquisa do Google que permite pesquisar trabalhos acadêmicos, literatura científica, periódicos de universidades e artigos variados relacionados à ciência. Foi lançado em Novembro de 2004 e passou a oferecer buscas em língua Portuguesa em 10 de janeiro de 2006 (TORRES-SALINAS; RUIZ-PÉREZ; DELGADO-LÓPES-CÓZAR, 2009).

Para extrair dados do Google Scholar, utilizou-se o "Harzing's Publish or Perish"², um *free software* que recupera e analisa citações acadêmicas e usa o Google Scholar para obter dados brutos, para então analisá-los e calcular uma série de métricas. Os resultados ficam disponíveis e podem ser copiados para planilhas ou serem salvos como um arquivo de texto (para futura referência ou análises adicionais). Dentre as análises e cálculos, pode-se obter (PUBLISH OR PERISH, 2014): Número total de documentos; Número total de citações; Número médio de citações por artigo; Número médio de citações por autor; Número de trabalhos por autor; Número médio de citações por ano; Duas variações de cada índice h; Análise do número de autores por artigo.

Além das diferentes estatísticas apresentadas, o aplicativo calcula as seguintes métricas de citação (MODESTO, 2010):

- **Hirsch's h-index**³: o índice baseia-se no conjunto dos trabalhos mais citados de um pesquisador e o número de citações recebidas de pessoas em outras publicações. O índice pode ser aplicado na análise da produtividade e do impacto de um grupo de cientistas, departamento, universidade ou país.
- **Egghe-g index**⁴: quantifica a produção científica dos físicos e outros cientistas com base no registro de suas publicações. O objetivo é melhorar o índice-h, dando mais peso aos artigos citados.
- **Contemporary h-index**⁵: Pretende também melhorar o índice-h, ao dar mais peso a artigos acadêmicos recentes de forma a incentivar a manutenção constante do nível de atividade de produção.
- **Zhang's e-index**⁶: O objetivo do índice é marcar a diferenciação entre os cientistas com os mesmos índices-h, mas com diferentes padrões de citação.
- **Age-weighted citation rate (AWCR) e AW-index**⁷: taxa de citação ponderada de idade que identifica o número médio de citações de um conjunto completo de trabalhos, ajustada a idade de cada trabalho individualmente. No aplicativo a execução do indicador difere da definição proposta pelo autor de somar todos os artigos ao invés de apenas de um núcleo-h de artigos.
- **Individual h-index (original)**⁸: O indicador divide o padrão índice-h pelo número médio de autores dos artigos que contribuem para o indicador a fim de reduzir os efeitos da coautoria.

² <http://www.harzing.com/pop.htm>

³ http://arxiv.org/PS_cache/physics/pdf/0508/0508025v5.pdf

⁴ <http://profs.sci.univr.it/~squassina/papers/Indexmetric/Egghe-number.pdf>

⁵ <http://arxiv.org/abs/cs.DL/0607066>

⁶ <http://www.plosone.org/article/info:doi/10.1371/journal.pone.0005429>

⁷ http://profs.sci.univr.it/~squassina/papers/Indexmetric/Jin_AR-index.pdf

⁸ <http://profs.sci.univr.it/~squassina/papers/Indexmetric/Compare-index.pdf>

- **H-index individual (variação PoP):** o software implementa uma alternativa ao individual h-index com uma abordagem diferente: ao invés de dividir o índice-h total, ele primeiro normaliza o número de citações para cada artigo, dividindo o número de citações pelo número de autores. Em seguida calcula o índice-h das citações normalizadas coletadas.
- **Multi-authored h-index⁹:** usa a contagem fracionada de artigo em vez da contagem reduzida de citação para artigos de autoria compartilhada, e então determina o índice hm demulti-autoria baseado em tabela efetiva resultante dos artigos utilizados na contagem das citações.

4 RELATO DE EXPERIÊNCIA

Atualmente a ECA-USP conta com 48 periódicos impressos (entre correntes e não correntes); 29 periódicos online¹⁰, sendo que destes, 12 fazem parte do Portal de Revistas da USP¹¹

A biblioteca da Escola de Comunicações e Artes da USP (ECA-USP), desde 2013 conta com um novo serviço de projetos, voltado para o desenvolvimento de produtos e serviços que atendam a novas demandas de Estudantes, Docentes e Pesquisadores. Este serviço, Com o objetivo de estabelecer um apoio mais próximo em relação às publicações da ECA, quanto às demandas por internacionalização, modernização (gestão por software) e profissionalização das atividades envolvidas na publicação de periódicos científicos (fluxo de: submissão, revisão por pares, normalização, publicação, etc) e também, no que se refere aos serviços do Programa de Apoio às Publicações Científicas Periódicas da USP¹², elaborou-se um formulário, enviado aos contatos das revistas, a fim de obter dados sobre o desenvolvimento e as necessidades destes periódicos.

Até o momento (Fevereiro de 2014), foram obtidos dados de 12 periódicos científicos. Destes, 9 participam do Programa de Apoio às Publicações Científicas Periódicas da USP e também fazem parte do Portal de Revistas da USP. São revistas heterogêneas abrangendo as áreas de: comunicação, turismo, artes cênicas, artes visuais, audiovisual.

Questionados se possuíam dados bibliométricos da Revista Científica, apenas 5 reponderam afirmativamente, porém não especificaram que tipo de fonte ou dados.

Neste primeiro momento (Fevereiro de 2014), realizou-se a busca dos títulos dos periódicos no software, sem limitação de datas, obtendo os dados e as análises geradas. Alguns dados

⁹ <http://iopscience.iop.org/1367-2630/10/4/040201/fulltext>

¹⁰ <http://www3.eca.usp.br/biblioteca/acervo/periodicosdaeca>

¹¹ <http://goo.gl/uZYO60>

¹² http://www.leginf.usp.br/?post_type=portaria&p=10532

preliminares: das 12 revistas a que obteve maior recuperação de *papers* é da área de comunicação, com 836 registros. O maior número de citação de um único artigo (210) data de 1995 e pertence a esta mesma revista. Outra publicação obteve o menor número de registro recuperado (15), e um artigo desta mesma publicação obteve 15 citações, mas o maior número de anos (33). Uma das revistas obteve 16 registros encontrados e nenhuma citação (0). Provavelmente por problemas de disponibilização dos metadados de uma das publicações, não obteve registro algum. O maior número de autores encontrado foi de 5 em um único artigo. Muitos outros dados bibliométricos podem ser analisados e servir de subsídio para a tomada de decisões dos editores científicos destas revistas.

4 CONSIDERAÇÕES FINAIS

Pode-se considerar que o Google Scholar é útil no nível micro, como uma ajuda para os autores, pesquisadores e editores na busca rápida, fácil e direta de documentos de texto completo, e em identificar as citações de seus trabalhos. Sobretudo é útil para a literatura não Anglo-Saxônica, que é menos controlada pelos sistemas dominantes de informação bibliométrica no mundo da ciência (WoS, SCOPUS), para as áreas que não utilizam preferencialmente as revistas científicas como meio de comunicação (ciências humanas, ciências sociais, etc) e para localizar citações de livros, teses, relatórios e artigos publicados em revistas secundárias não constituídas na chamada "ciência *mainstream*" (TORRES-SALINAS; RUIZ-PÉREZ; DELGADO-LÓPES-CÓZAR, 2009).

Assim, a manutenção dos conjuntos de dados bibliométricos de pesquisa requer o conhecimento do comportamento científico da área e conhecimento tecnológico avançado, a fim de extrair, analisar, organizar e armazenar os dados para que possam ser devidamente preservados e para que outros pesquisadores sejam capazes de consultar de forma eficaz a informação, sobretudo para a geração de métricas importantes para a tomada de decisão. Desta forma, entende-se que, de acordo com todo o exposto, o desenvolvimento de serviços de dados bibliométricos nada mais é do que a **extensão natural do papel tradicional das bibliotecas acadêmicas** (NIELSEN; HJØRLAND, 2014, grifo do autor).

REFERÊNCIAS

GARCIA, Rodrigo Moreira. **Modelos de comportamento de busca de informação:** contribuições para a Organização da Informação. 2007. 139f. Dissertação – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista – UNESP, Marília. 2007. Disponível em: <<http://eprints.rclis.org/15386/>>. Acesso em: 23 Mar. 2014.

MODESTO, Fernando. **Publicar ou Perecer:** software para a biblioteca avivar o pesquisador. 2010. Disponível em: <http://www.ofaj.com.br/colunas_conteudo.php?cod=540>. Acesso em: 26 fev. 2014.

NIELSEN, Hans Jørn; HJØRLAND, Birger. Curating research data: the potential roles of libraries and information professionals. **Journal of Documentation**, v.70, n.2, p.221-240, 2014
Disponível em: <<http://dx.doi.org/10.1108/JD-03-2013-0034>>. Acesso em: 22 fev. 2014.

PUBLISH OR PERISH. **About.** 2014. Disponível em:
<<http://www.harzing.com/pop.htm#about>>. Acesso em: 27 fev. 2014.

REBIUN; CRUE. **III Plan Estratégico de Rebiun 2020.** 2011. Disponível em:
<http://www.mcu.es/bibliotecas/docs/MC/ConsejoCb/CTC/Bib_Univ/Planestrategico2020.pdf>. Acesso em: 20 dez. 2013.

TORRES-SALINAS, Daniel; RUIZ-PÉREZ, Rafael; DELGADO-LÓPES-CÓZAR, Emilio. Google Scholar como herramienta para la evaluación científica. **El profesional de la información**, v. 18, n.5, 2009. Disponível em:
<<http://ec3.ugr.es/publicaciones/d700h04j123154rr.pdf>>. Acesso em: 23 fev. 2014.